Check for updates

Insulator defect detection based on improved Yolov5s

Dehong Wei[1], Bo Hu[1]*, Chaoyang Shan[1] and Hanlin Liu[2]

[1]Surveying and Mapping Engineering Faculty, Guangdong University of Technology, Guangzhou, China,
[2]School of Computer Science, Guangdong University of Technology, Guangzhou, China

The stable operation of a power supply system is inseparable from the work of detecting defects in transmission lines. However, the insulator defect detection model based on deep learning is widely used in wire inspection work. Therefore, this paper proposes an improved YOLOv5s insulator defect detection model in order to solve the problems of insufficient training data and low recognition accuracy of the target detection model in the real-time detection of small target insulator defects. To expand and enhance the training data, experiments were conducted using the addition of noise and random black blocks. The spatial and channel weight coefficients were obtained by adding an attention mechanism (Convolutional Block Attention Module, CBAM), and the dimensions of the input feature maps were transformed to enhance the model's ability to extract and fuse small target defect features. Experiments show that with Faster RCNN, YOLOv3, SSD and YOLOv4 comparison experiments verified that the algorithm achieves 97.38% detection accuracy for insulators and 93.32% detection accuracy for small target insulator defects with a fast detection speed, which is a better solution to the problem of detecting insulator defects with too small a proportion in the image.

KEYWORDS

insulators, YOLOv5S, defect detection, cbam attention mechanism, deep learning

1 Introduction

The normal operation of modern society is dependent on a stable power supply system. Insulators, as one of the most commonly used components in power line equipment, are widely used in the conductor insulation and cable support of power transmission systems (Zhao et al., 2021a). However, traditional power inspection method mainly relies on manual inspection by manually climbing to high places to conduct visual inspections. This not only wastes a lot of time and energy, but also makes it difficult to ensure the safety of inspection personnel, and the quality of inspections is easily affected by weather factors. Insula-tors are located in a natural outdoor environment and are prone to bursting and flashing faults under the erosion of rain and Sun (HUANG et al., 2022), causing serious damage to the reliable operation of the transmission system; hence the inspection of power lines should be carried out regularly and comprehensively. However, insulators are often erected in the air and are subject to the complexity of the line background and air angle sight distance restrictions. Moreover, traditional manual inspection methods are difficult, time-consuming, and labor-intensive, and there are certain operational safety risks (Chen et al., 2019). Therefore, there is even a need to rely on ultrasonic, infrared (DENG et al., 2019; ZHAO et al., 2021b), and other professional fault detection equipment for accurate discrimination, and it is difficult to achieve a widespread efficient inspection.

Computer vision techniques represented by deep learning have emerged in recent years and are now widely used in the detection of transmission line faults (Zhao et al., 2020a; Wen

et al., 2021). Many scholars have used target detection algorithms in deep learning to detect and identify defects in insulators in transmission lines, among which the two-stage target detection algorithm represented by Faster R-CNN (Faster Region–Convolutional Neural Network) (Zhao et al., 2019a; Li et al., 2020) and the single-stage target detection algorithm represented by YOLO (You Only Look Once) (Liu et al., 2021b; Gao et al., 2022) and SSD (Single Shot MultiBox Detector) (YANG et al., 2020; LIU et al., 2021a) are the most popular in line fault detection research. In the literature (Lei and Sui, 2019), a deep convolutional neural network model based on Faster R-CNN is proposed to transform the target classification problem into a target detection and recognition problem, using ResNet-101 (Residual Neural Network-101) pre-trained weights to locate insulators and bird nests effectively, with the average accuracy value of the model reaching 97.6%. The Literature (Zhao et al., 2019b) introduces skip connections and reduces the number of network layers by adjusting the order of ac to enhance Faster R-CNN's detection of small targets. The literature (Jiyu et al., 2021) introduces multiscale training and employs adversarial networks to generate defective image occlusion region features which enable the detection network to analyze difficult occlusion situations in greater depth, thereby improving the sample detection rate.

However, the training process of the two-stage target detection model includes a region proposal network, which makes the model structure more complex and has a larger number of parameters, making it difficult to quickly produce target detection results. In contrast, single-stage target detection models such as YOLO and SSD can directly perform classification and regression after completing backbone feature extraction, significantly improving detection speed. This has resulted in one-stage target detection models becoming the current research hotspot. The literature (ZHAO et al., 2020b) addresses the problem of inaccurate target detection due to overlapping and intersecting target frames of the fixtures. It uses the similarity of intersecting regions as the contextual information of the targets in order to design the occlusion relationship module, which is embedded into the SSD model for the target detection of mutually occluding features. The literature (Liu et al., 2020) proposes an improved SSD algorithm using the lightweight network MobileNet as the feature extraction network and two multiscale feature fusion methods, Channel Plus and Channel Concatenate, to improve the accuracy and speed of the algorithm.

The YOLO family of algorithms has been widely used and iterated in many versions in the field of target detection over the years (Redmon et al., 2016). In the literature (WU et al., 2019), Crop-MobileNet is used as a feature extraction network for the YOLOv3 model to meet the requirements of fast UAV inspection. The stability of the prior frame was improved using the Euclidean distance based K-means++ algorithm, which also improved the detection speed significantly. In (Zhang et al., 2020), an improved Dense-YOLOv3 algorithm is proposed in order to design a dense network (Dense-Net) instead of the single convolutional layer in the original network, thereby achieving multilayer feature multiplexing and the fusion of insulators, and improving the detection accuracy. The literature (Gao et al., 2021) introduced the Triplet Attention module and added the Dense Block module into the YOLOV5-based backbone network to alleviate the

problem of gradient disappearance. The results show that the accuracy of the improved network structure reaches 94.5%. The literature (Chen et al., 2023) proposes an improved insulator defect identification algorithm Insu-YOLO based on the latest YOLOv8 network. It adopts a lightweight content-aware feature recombination (CARAFE) structure and has an accuracy of 95.9%, which is 3.95% higher than the YOLOv8n baseline model. The literature (Tang, 2021) combines U-Net and YOLOv5 for the segmentation and localization of insulators. It also introduces a residual structure to reduce the effect of gradient disappearance, introduces an attention mechanism to correct the feature weights, and cuts high-resolution images before detection, improving the effectiveness and efficiency of the model. However, since the percentage of insulator defects in the image is too small, the common target detection model does not obtain good results in small feature extraction and detection, which makes it difficult to play a significant role in practical engineering.

There are already many topics related to the detection of power line components using deep learning technology, but there are still many problems. On the one hand, the main detection objects are single. The detection objects are basically various insulators, and other power There are few studies on circuit components, and on the other hand, there are almost no studies on experimental subjects with small samples and small targets.

Therefore, to improve the detection accuracy of small target insulator defects, this paper proposes an enhanced YOLOv5s insulator defect detection model by adding an attention mechanism and adjusting the feature map size. First, to compensate for the lack of training data, the training dataset is expanded by data augmentation through noise enhancement and the random addition of black blocks, data pasting and copying. Using this as a basis, the network structure is improved by adding the attention mechanism CBAM. Additionally, the size of the feature map is adjusted to enhance the feature extraction and fusion capability of the model for insulator defects, which improves the network detection accuracy with little impact on the detection speed. In comparison tests, the mAP (mean average precision) of the YOLOv5s model after adding the attention mechanism was 93.65%, an improvement of 1.32%. After readjusting the feature map size, the mAP was improved by another 1.70%–95.35%, and there was only a small decrease in detection speed which did not affect the detection efficiency. The comparison experiments verify that the algorithm in this paper is a better solution to the problem of detecting insulator defects with too small a proportion in the image.

2 Materials and methods

2.1 YOLOv5s model structure

YOLOv5 is the latest detection model developed by the YOLO (You Only Look Once) target detection network algorithm series after years of iterative updates. YOLOv5 model results are easy to implement, can improve target recognition and detection accuracy, and are widely used in Insulator defect detection (Luo et al., 2021; Ding et al., 2022; Li et al., 2022; Zhang et al., 2023). Due to its fast detection speed and high level of accuracy, it is widely used in various target detection fields. The basic architecture of the
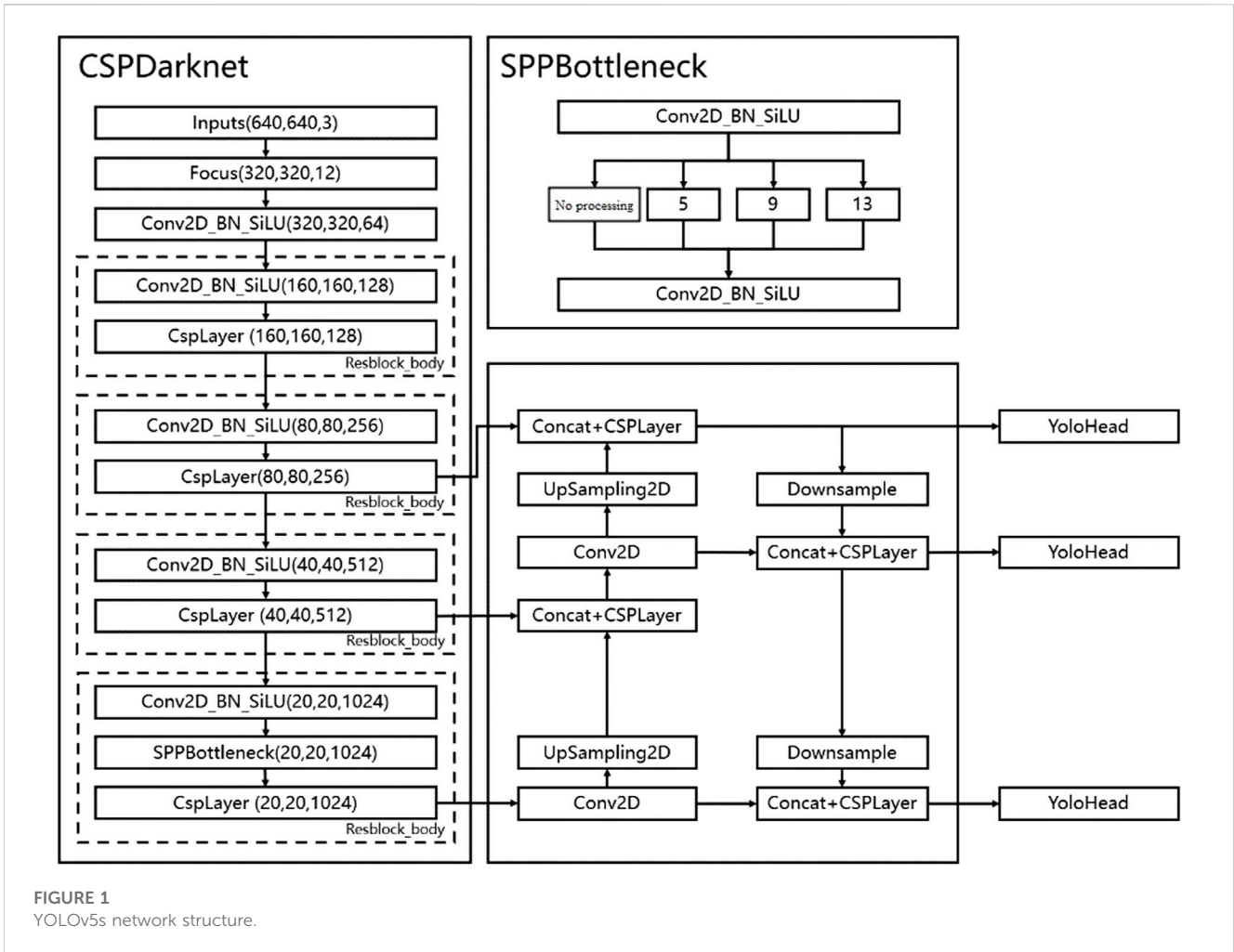
FIGURE 1
YOLOv5s network structure.

YOLOv5 network for different requirements, including small (s), medium (m), large (l), and extra large (x) models, is the same, except that the depth and width of the network is adjusted in each sub-module in different versions. YOLOv5s is a variant of YOLOv5 with sub-modules of smaller depth and width dedicated to smaller images and to obtaining faster detection speeds, which is important for applications in real-time systems or systems with low latency requirements. In addition, YOLOv5s can be used in combination with other models in many different scenes, easily scaling to larger images and more complex scenes. The structure of the YOLOv5s network is shown in Figure 1. It consists of three main parts: a backbone feature extraction network (Backbone), an enhanced feature fusion network (FPN), and a head detection network (Head).

As the feature extractor of the detection system, the structure of the backbone feature extraction network, also called CSP-Darknet, consists of multiple underlying convolutional and residual blocks. The residual structure enables the fusion of shallow and deep image features. By stacking multiple layers of residual and structural blocks, the network is able to learn more complex feature representations while reducing the gradient disappearance problem and not losing image features due to the increase of convolutional network depth. The Focus module and the Spatial Pyramid (SPP) module are also used in CSP-Darknet to
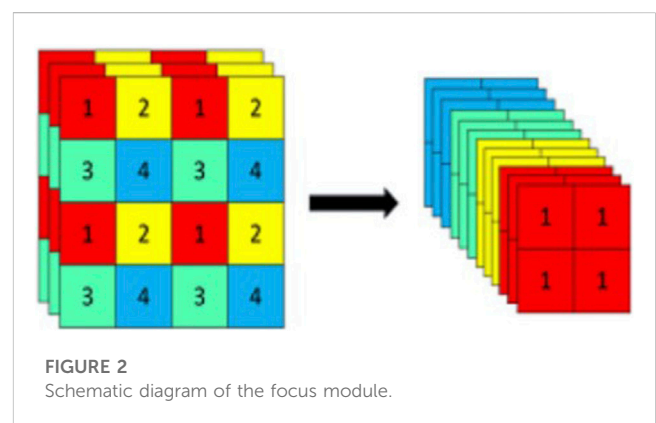

FIGURE 2
Schematic diagram of the focus module.

downsample the feature maps. The Focus module divides each 2×2 pixel in the image into a small block, and extracts and stitches the pixels at the same position in each block into a new independent feature layer before stacking to increase the number of feature layer channels, as shown in Figure 2.

Through the processing of the Focus module, the image matrix size is reduced to one-half of the original size, and the number of channels is changed from three channels to twelve channels. This increases the number of data channels and reduces the spatial

FIGURE 3
SiLU activation function curve.
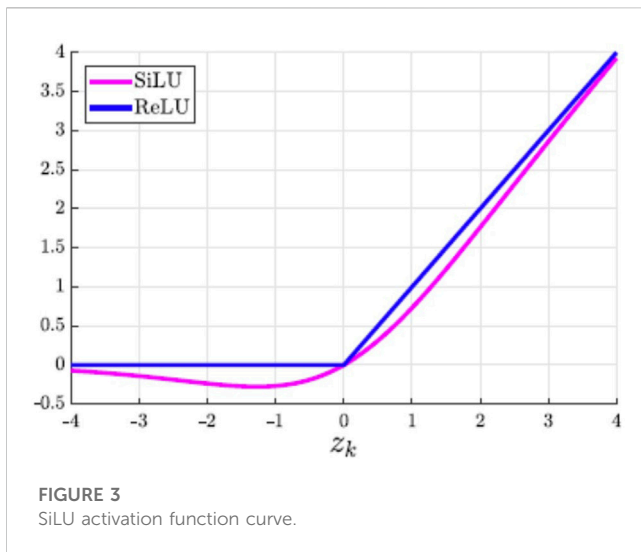
dimension of the image without losing the feature information. The SPP module performs feature extraction on the convolved feature matrix using maximum pooling with different pooling kernel sizes in the last residual structure block, which can greatly improve the perceptual field of the feature map and can extract the most significant target features.

FPN (Feature Pyramid Network) layers are often used to extract multiple layers of features from an image, and these feature layers are combined together to form a pyramid-like structure containing both upsampling and down sampling. This pyramid-like structure is used to extract features at different scales from the three effective feature layers output by the backbone feature extraction network and subsequently fuse them in a stacked fashion to retain the effective feature information of targets at different scales, thus helping the algorithm to better identify targets at different scales. With the backbone feature extraction network CSP-Darknet and the feature fusion network FPN, we have been able to obtain three enhanced feature layers. The YoloHead detection network is responsible for processing the extracted feature layers and converting them into the output of target detection, including the category, location, and bounding box of each target. The YoloHead is divided into three sizes according to the size of the feature map. Different detection modules with dimensions of (80×80×number of categories), (40×40×number of categories) and (20×20×number of categories) are used to detect large, medium and small targets respectively, and to predict the bounding box coordinates, class labels, and confidence scores of the objects in the image. After the final prediction results are obtained, score ranking and non-great suppression screening are also performed to draw the detection frame with the best results.

2.2 SiLU activation function

Using activation functions on the convolution results can enhance the nonlinear feature representation of the model.

SiLU (also known as Swish), represented in Figure 3, is a recently introduced activation function adopted as an alternative to the

popular ReLU (Rectified Linear Unit) activation function. It is defined as:

$$f(x) = x * sigmoid(x) \tag{1}$$

In the above, sigmoid is the standard sigmoid activation function:

$$sigmoid(x) = 1 / (1 + exp(-x)) \tag{2}$$

SiLU, similar to RELU, is a segmented linear activation function that can be used to back-propagate training parameters for neural networks. However, compared to ReLU and other activation functions, SiLU has some key differences: non-monotonicity, smooth derivability, and saturation. Firstly, unlike the monotonically increasing function ReLU, SiLU is non-monotonic, which means it can have both positive and negative slopes, allowing for the capture of both positive and negative correlation information in the data. Secondly, SiLU is smooth and continuously differentiable compared to ReLU with discontinuous derivatives at the origin, which makes it easier to optimize the neural network parameters using gradient descent-based optimization methods. Additionally, the bounded output range (−1, 1) for SiLU helps prevent the output of the neuron from becoming too large or too small. Thirdly, like ReLU, SiLU has saturation behavior, which means it flattens out when the input is too large or too small. However, the saturation behavior of SiLU is not as severe, and it is less prone to gradient disappearance problems. SiLU is a more effective alternative to other activation functions, and it has been shown to perform well in various deep learning models such as image processing and natural language processing. It is still relatively new, however, and more research is needed to fully understand its characteristics and potential benefits.

2.3 Improvements to the YOLOv5s model

The CBAM (Convolutional Block Attention Module) allows convolutional neural network models to focus on a specific part of the input rather than processing the entire input equally. This is particularly effective for tasks such as target identification. The CBAM consists of two parts: a channel attention module and a spatial attention module (Figure 4). The channel attention module learns which channel feature maps are most important for the task and usually consists of a fully connected layer and a sigmoid activation function that establishes correlations between channels, obtains the weight coefficients for each channel, weights them onto the original features, and completes the recalibration of the original features in the channel dimension. The input in the spatial attention module is a feature map (h×w×c) that has been passed through the channel attention module, and all input channels are pooled into two in the channel dimension to obtain two (h×w×1) feature maps, followed by a convolution kernel of size 7×7, which is convolved to form a (h×w×1) feature map. Finally, the spatial feature weights are obtained by the activation function to generate a weighted feature map reflecting the importance of each channel and each pixel in the image, and then multiplied with the new feature map to obtain a feature map adjusted by the dual attention mechanism.
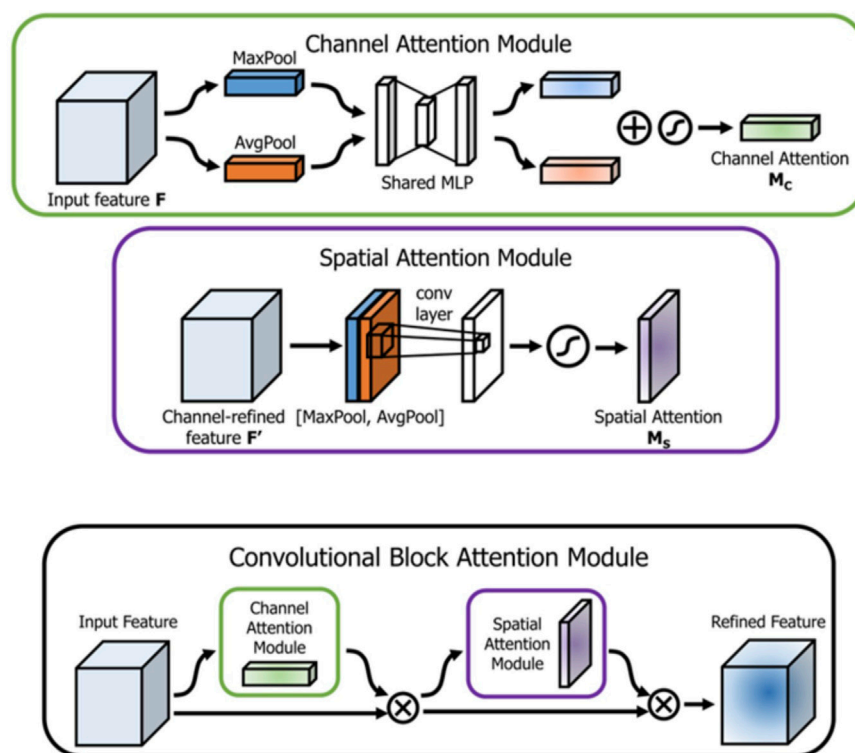
FIGURE 4
CBAM attention module.

To continue enhancing the detection of small target insulator defects, a shallower feature map is selected as one of the inputs to the feature fusion network, as shown in Figure 5.

Because insulator defects occupy so few pixels, shallow feature maps with small sensory fields are suitable for detecting small-scale objects. The feature maps produced by the first few layers of the network will have more prominent and easily distinguishable small features, and the model can more easily capture the rough location and rough shape of these small targets before using deeper feature maps to further refine the location and shape information. In addition, shallow feature maps require fewer parameters and require less computation than deep feature maps, which can make the model faster and more efficient. This is especially important for real-time target detection applications, where speed is a key factor. Therefore, feature maps with a feature map size of (80, 80, 256) are discarded and a shallower feature map (160, 160, 128) is used for feature extraction.

3 Experimental results and data analysis

3.1 Dataset processing

The training image dataset used by the model is from the Chinese Power Line Insulator Dataset (CPLID). With 748 images, the original dataset is small and lacks sufficient diversity. Since deep learning networks require a large amount of rich training data to obtain better learning results, data augmentation is performed using a number of methods such as enhanced noise and randomly adding black blocks in order to make the trained model more robust. Adding noise to the data is one way to enhance the data. A common method is to add Gaussian noise, which is a type of noise that obeys a normal distribution. It is also possible to add 'pretzel' noise, which consists of random black and white pixels scattered throughout the image. The randomly selected pixel value can be set to zero (black) or 255 (white). Randomly adding black blocks can simulate the situation where the actual target is obscured and helps to improve the recognition accuracy of the obscured target. Adding noise to the data can help improve the generalization of the model, but it can also make the data harder to process and may negatively affect performance if the noise is too strong. The effect of partial image enhancement is shown in Figure 6.

After data enhancement, the dataset has 2000 images, which are divided into training, validation and test sets in the ratio of 7:2:1.

3.2 Experimental environment

The operating system used in the experimental platform is Windows 10, the CPU is Intel Core i7-10700 KF @ 3.80GHz, the RAM is 64GB, and the GPU is NVIDIA GeForce RTX 3080 with 8 GB of video memory.

3.3 Model training

The training image size is uniformly set to (640×640). The model is trained for 100 epochs, the batch size is set to 8, and the

FIGURE 5
Shallow feature map as input.

FIGURE 6
Data enhancement effect.
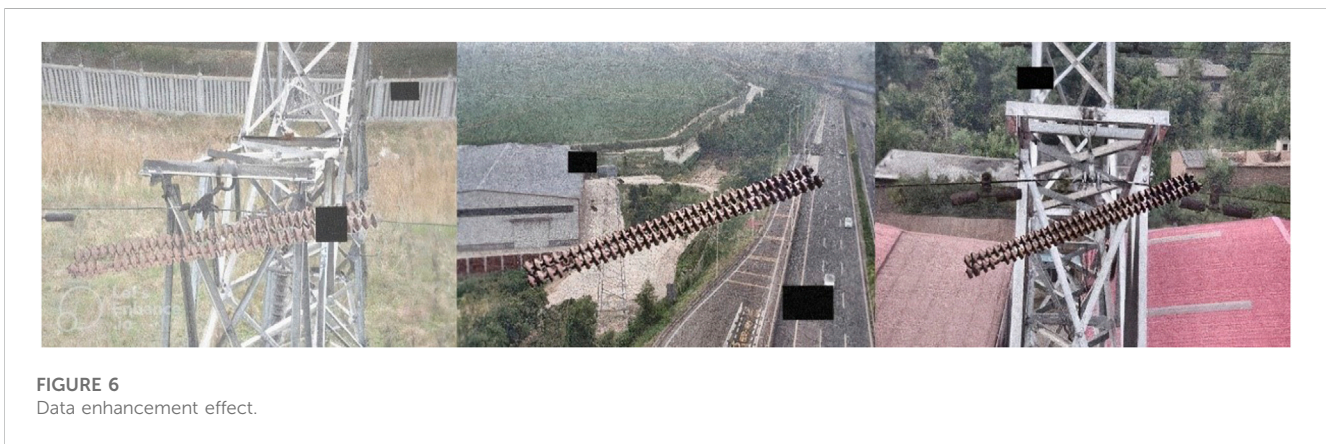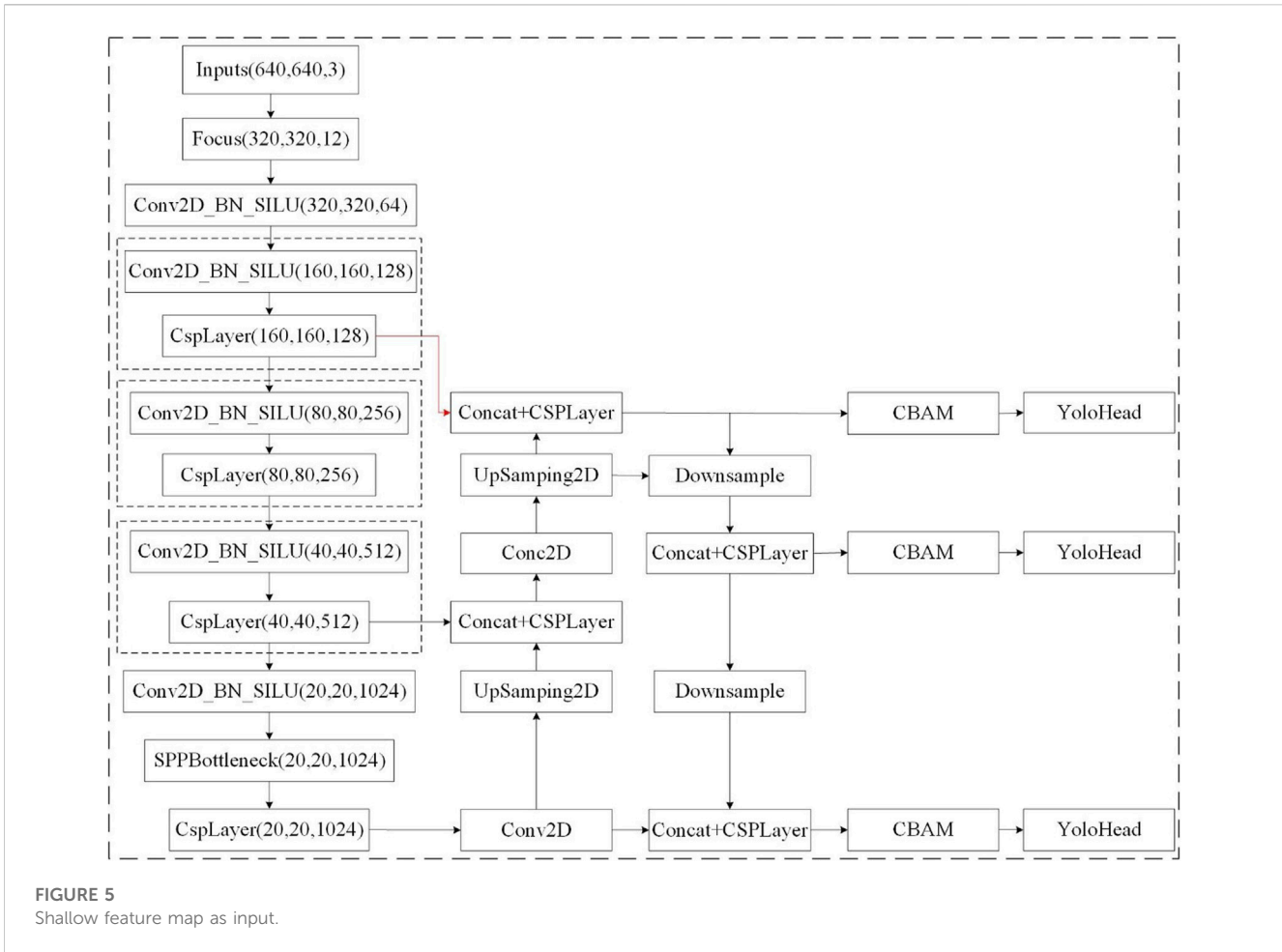
initial learning rate is set to 1e−3. The curve in Figure 7 shows the changes in training loss.

3.4 Evaluation indicators

The evaluation metrics used to verify the accuracy in this paper are defect detection accuracy, insulator detection accuracy, mean average precision (mAP), which is the sum of the average accuracy of all

categories divided by the number of categories, and the number of frames per second (FPS) that the model can detect. The average precision (AP) is calculated as follows, where the AP is for a certain class of samples:

$$AP = \frac{1}{m}\sum_{i}^{m} P_i = \frac{1}{m}{}^*P_1 + \frac{1}{m}{}^*P_2 + \cdots + \frac{1}{m}{}^*P_m \qquad (3)$$

In Eq. 3, the sample has m positive cases, where each positive case corresponds to a recall. The maximum accuracy p is calculated
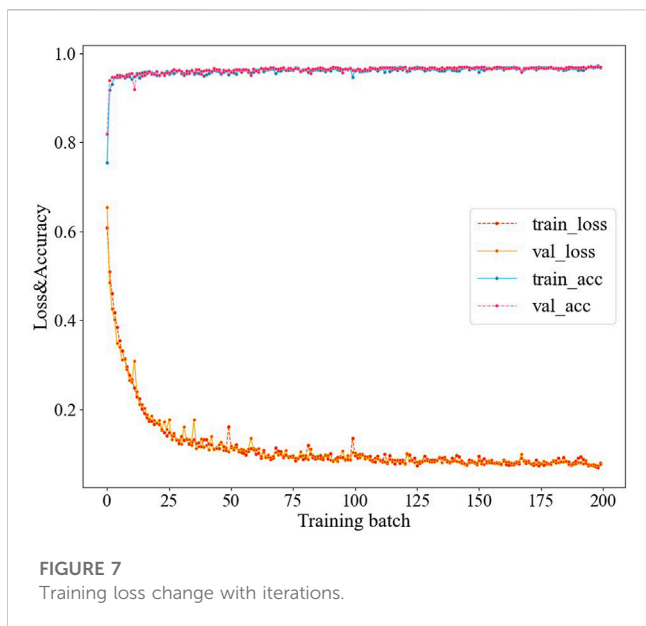
FIGURE 7
Training loss change with iterations.

TABLE 1 Comparison of target detection model performance.

Model	Defect	Insulator	mAP	FPS
Faster RCNN	86.33	89.53	87.93	20.33
YOLOv3	83.65	86.91	85.28	43.98
SSD	88.68	89.28	88.98	42.77
YOLOv4	90.97	91.39	91.18	47.33
YOLOv5s	91.43	93.21	92.33	48.68

for each m, and the average value of each p is calculated to obtain the AP of the class. The mAP is obtained by averaging the APs of all the C categories contained in the dataset, as follows:

$$mAP = \frac{1}{C}\sum_{j}^{C} AP_j \qquad (4)$$

3.5 Experimental results and analysis

Table 1 compares the detection performance of different target detection models with the YOLO series of models. The detection accuracy of Faster RCNN for insulators and defects is higher than YOLOv3, but it is a two-stage model with a more complex structure. Consequently, detection takes longer, and the FPS is only 20.33, which is less than half the detection speed of SSD and the YOLO series of models. The SSD model has a somewhat higher detection accuracy than YOLOv3, but after iterations of the YOLO series the mature YOLOv4 and YOLOv5s have improved even more in terms of detection accuracy and speed, surpassing the SSD model, especially YOLOv5s. The SiLU activation function shows its robustness on YOLOv5s with a 91.43% insulator defect detection

accuracy, a 92.33% mAP, and the fastest detection rate of 48.68 frames per second.

In order to verify the effect of the YOLOv5s improvements, we tested the performance of the YOLOv5s model before improvement, tested the performance after adding the attention mechanism, and then tested the performance after changing the input feature map size. The results of the comparison are shown in Table 2. The CBAM module enables the model to learn a greater number of accurate spatial and channel features without losing image features due to the increase in network depth, and increases the insulator detection accuracy to 95.07%. In addition, after transforming the input feature map and switching to shallow features, the detection accuracy of small target defects was further increased to 93.32%, while the insulator detection accuracy also increased and the final mAP was improved. However, due to the addition of the attention mechanism, the more complex model structure has an impact on the detection rate, which drops from 48.68 frames per second to about 43 frames per second. Although there is a small decrease in detection speed, it can still meet the demand for rapid detection in real-time.

Figure 8 shows an example of the detection results of the YOLOv5s model before and after the improvements. Figure 8B shows that after adding the attention mechanism and transforming the size of the input feature map, the model is able to detect accurately insulator targets of all sizes. Moreover, the detection frame is more refined, maintaining a high confidence level of recognition and improving the confidence level for the detection of small target defects.

4 Discussion

4.1 Model performance comparison analysis

The experimental results presented in Table 1 highlight the trade-offs among var-ious target detection models, particularly focusing on the comparison with the YOLO series. It is evident that Faster RCNN achieves higher detection accuracy for insulators and defects compared to YOLOv3. However, its two-stage model structure contributes to increased complexity and longer detection times, resulting in a lower frames-per-second (FPS) rate of 20.33.

On the other hand, the single-stage models, SSD and the YOLO series (v3, v4, and v5s), exhibit superior detection speed. The YOLO series, especially the mature YOLOv4 and YOLOv5s, surpass both Faster RCNN and SSD in terms of both detection accuracy and speed. YOLOv5s, utilizing the SiLU activation function, demonstrates remarkable robustness, achieving a 91.43% accuracy for insulator defect detection, a mean average precision (mAP) of 92.33%, and the fastest detection rate at 48.68 frames per second. By incorporating the CBAM module, the model gains the ability to learn more accurate spatial and channel features, leading to a notable increase in insulator detection accu-racy to 95.07%. Although the introduction of the attention mechanism impacts the de-tection rate, causing a decrease from 48.68 to approximately 43 frames per second, the model still meets the requirements for real-time detection.

TABLE 2 Performance comparison of YOLOv5s model before and after improvements.

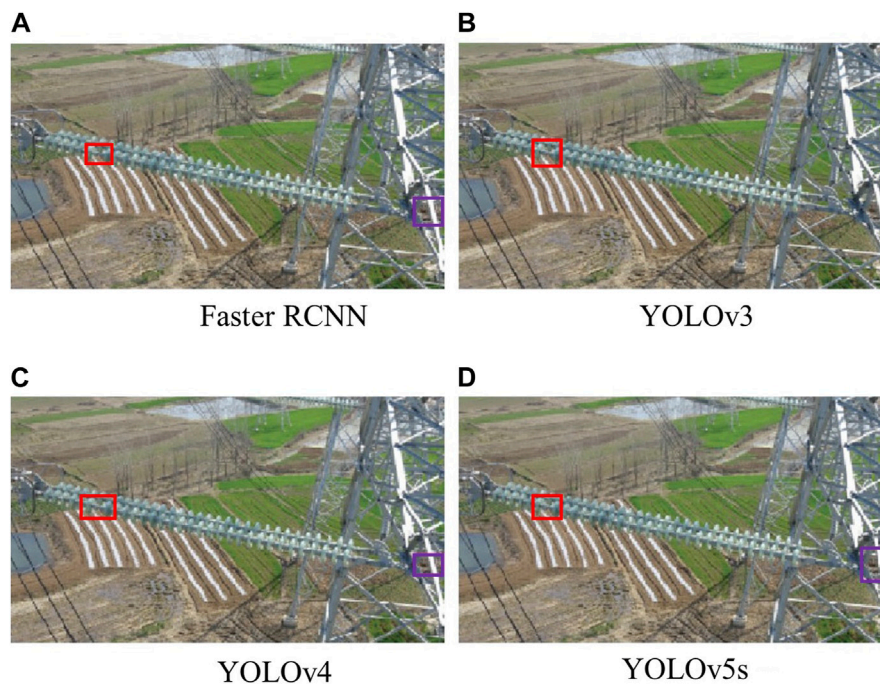| Model	Defect	Insulator	mAP	FPS
Yolov5s	91.43	93.21	92.33	48.68
Yolov5s with CBAM	92.23	95.07	93.65	43.30
Yolov5s with CBAM and transformation of the input characteristics map	93.32	97.38	95.35	42.60

FIGURE 8
(A) Faster RCNN; (B) YOLOv3; (C) YOLOv4; (D) YOLOv5.

4.2 Data validity

The utilization of data augmentation techniques, such as introducing noise and random black blocks, proves instrumental in bolstering the training data for object detection models. This strategy enhances the model's robustness to variability, fosters improved generalization to diverse scenarios, and promotes the extraction of relevant features critical for accurate object detection. Moreover, data augmentation addresses challenges associated with limited datasets, mitigates overfitting, and enhances the model's adaptability to real-world variations. Overall, augmenting the dataset with noise and random black blocks serves as a powerful tool to elevate the detection accu-racy and performance of object detection models.

4.3 Future prospects and limitations

Given the limitations imposed by the available sample dataset, our research has thus far concentrated on the application of deep learning object detection algorithms to address insulator defect

detection. Looking ahead, we envision an expansion of our efforts to explore and develop additional defect detection methods. Our goal is to build upon the existing foundation, refining and innovating techniques that go beyond ob-ject detection. This future research aims to elevate the automation levels of power line inspections, contributing to more comprehensive and effective methods for enhancing the reliability and efficiency of defect detection in the field of electrical power infra-structure.

5 Conclusion

This paper has proposed an improved YOLOv5s model to resolve the problem of low accuracy of current target detection models for small target insulator defects. Experiments were conducted to compare its performance with those of Faster RCNN, YOLOv3, and YOLOv4. The results show that by adding the CBAM attention mechanism to the last layer of the YOLOv5s enhanced feature fusion network, more effective insulator features can be extracted by increasing the feature weights of space and channels. The detection of small target

insulator defects can be enhanced by adjusting the size of the input feature map, with the final mAP for insulator defect detection reaching 93.32%, and the mAP for insulator detection reaching 97.38%. The results verify the effectiveness of the improved model in increasing detection accuracy. While the detection speed showed a small decrease compared to the original YOLOv5s model, the decrease did not affect the detection efficiency in real-time. The main focus in subsequent research is network model miniaturization in order to improve the detection speed while ensuring that the detection accuracy is maintained. Taken together, these findings further confirm the importance of this research, which will help advance the development of related research and applications.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Author contributions

DW: Conceptualization, Formal Analysis, Methodology, Project administration, Resources, Writing–original draft. BH: Supervision, Validation, Writing–review and editing. CS: Data curation, Software, Visualization, Writing–review and editing. HL: Data curation, Writing–review and editing.

Funding

Acknowledgments

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Chen, H., He, Z., Shi, B., and Zhong, T. Research on recognition method of electrical components based on YOLO V3[J]. *IEEE Access*, 2019, 7: 157818–157829. doi:10.1109/access.2019.2950053

Chen, Y., Liu, H., Chen, J., Hu, J., and Zheng, E. (2023). Insu-YOLO: an insulator defect detection algorithm based on multiscale feature fusion. *J. Electron.* 12 (15), 3210. doi:10.3390/electronics12153210

Deng, H. L., He, Z. F., and Chen, L. Inspection of debonding defects of composite insulators by ultrasonic guide wave [J]. *High. Volt. Eng.*,2019,45(01):196–202. doi:10.13336/j.1003-6520.hve.20181229020

Ding, J., Cao, H., Ding, X., and An, C. High accuracy real-time insulator string defect detection method based on improved YOLOv5[J]. *Front. Energy Res.*, 2022, 10: 928164, doi:10.3389/fenrg.2022.928164

Gao, J., Chen, X., and Lin, D. (2021). Insulator defect detection based on improved YOLOv5[C]//2021 5th asian conference on artificial intelligence technology (ACAIT). *IEEE*, 53–58.

Gao, J. C., Zhang, J. H., Li, Y. N., and Li, C. Insulator burst fault identification based on YOLOv4 [J]. *Laser & Optoelectron. Prog.*,2022,59(02):130–137.

Huang, X. B., Gao, Y. H., Zhang, Y., Zhao, L., Wu, Y. Q., and Sun, S. Z. Glass insulator target recognition algorithm based on joint component grayscale algorithm and deep learning [J]. *Electr. Power Autom. Equip.*,2022,42(04): 203–209. doi:10.16081/j.epae.202201032

Jiyu, Y. I., Cifa, C., and Guoqiang, G. (2021). Aerial insulator detection of transmission line based on improved Faster RCNN. *Comput. Eng.* 47, 292.

Lei, X., and Sui, Z. (2019). Intelligent fault detection of high voltage line based on the Faster R-CNN. *J. Meas.* 138, 379–385. doi:10.1016/j.measurement.2019.01.072

Li, F., Xin, J., Chen, T., Wei, Z., Li, Y., et al. An automatic detection method of bird's nest on transmission line tower based on faster_RCNN[J]. *IEEE Access*, 2020, 8: 164214–164221. doi:10.1109/access.2020.3022419

Li, Y., Ni, M., and Lu, Y. Insulator defect detection for power grid based on light correction enhancement and YOLOv5 model[J]. *Energy Rep.*, 2022, 8: 807–814. doi:10.1016/j.egyr.2022.08.027

Liu, C., Wu, Y., Liu, J., and Sun, Z. Improved YOLOV3 network for insulator detection in aerial images with diverse background interference[J]. *Electronics*, 2021a, 10(7): 771, doi:10.3390/electronics10070771

Liu, X., Li, Y., Shuang, F., Gao, F., Zhou, X., and Chen, X. ISSD: improved SSD for insulator and spacer online detection based on UAV system[J]. *Sensors*, 2020, 20(23): 6961, doi:10.3390/s20236961

Liu, X. Y., Miao, X. R., Zhuang, S. B., Jiang, H., and Chen, J. Insulator detection based on lightweight deep convolutional neural network [J]. *J. Fuzhou Univ. Sci. Ed.*,2021b,49(02):196–202.

Luo, P., Wang, B., and Ma, H. Defect recognition method with low false negative rate based on combined target detection framework[J]. *High. Volt. Eng.*, 2021, 47: 454–464.

Redmon, J., Divvala, S., Girshick, R., et al. You only look once: unified, real-time object detection[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 05-June-2018, China, IEEE, 2016: 779–788.

Tang, X. Y., **ong H, Huang R, et al. Insulator mask acquisition and defect detection based on improved U-Net and YOLOv5[J]. *J. Data Collect. Process.*, 2021, 36(05): 1041–1049.

Wen, Q., Luo, Z., Chen, R., Yang, Y., and Li, G. Deep learning approaches on defect detection in high resolution aerial images of insulators[J]. *Sensors*, 2021, 21(4): 1033, doi:10.3390/s21041033

Wu, T., Wang, W., Li, Y. U., et al. Insulator defect detection method for lightweight YOLOV3 [J]. *Comput. Eng.*,2019,45(08):275–280. doi:10.19678/j.issn.1000-3428.0053695

Yang, G., Sun, C. W., Wang, D. W., Jin, T., Xu, C. Y., Lu, Z. B., et al. Comparative study of transmission line component detection models based on UAV front end and SSD algorithm [J]. *J. Taiyuan Univ. Technol.*,2020,51(02):212–219. doi:10.16355/j.cnki.issn1007-9432tyut.2020.02.009

Zhang, H., Li, J., and Zhang, B. (2020). Foreign object detection on insulators based on improved YOLO v3[J.]. *Electr. Power* 53, 49–55.

Zhang, T., Zhang, Y., Xin, M., Liao, J., and Xie, Q., A light-weight network for small insulator and defect detection using UAV imaging based on improved YOLOv5[J]. *Sensors*, 2023, 23(11): 5249, doi:10.3390/s23115249

Zhao, L. H., Gao, Q., Li, D. H., and Yu, X. Multiple insulators extraction method based on complexinfrared images [J]. *Laser J.*,2021a,42(05):62–67. doi:10.14016/j.cnki.jgzz.2021.05.062

Zhao, L. I. U., Zhang, L., Geng, M., et al. Object detection of high-voltage cable based on improved Faster R-CNN[J]. *CAAI Trans. intelligent Syst.*, 2019b, 14(4): 627–634.

Zhao, Z., Qi, H., Fan, X., Xu, G., Qi, Y., Zhai, Y., et al. Image representation method based on relative layer entropy for insulator recognition[J]. *Entropy*, 2020a, 22(4): 419, doi:10.3390/e22040419

Zhao, Z., Zhen, Z., Zhang, L., Qi, Y., Kong, Y., and Zhang, K. Insulator detection method in inspection image based on improved faster R-CNN[J]. *Energies*, 2019a, 12(7): 1204, doi:10.3390/en12071204

Zhao, Z. B., Jiang, Z. G., Li, Y. X., Qi, Y. C., Zhai, Y. J., Zhao, W. Q., et al. (2021b). Overview of visual defect detection of transmission line components. *J. Image Graph.* 26 (11), 2545–2560.

Zhao, Z., Jiang, A., Qi, Y., et al. Fittings detection in transmission line images with SSD model embedded occlusion relation module[J]. *CAAI Trans. intelligent Syst.*, 2020b, 15(4): 656–662.